EFFICIENT SYNTHESIS OF APPROXIMATELY CONSISTENT GRAPHS FOR ACOUSTIC MULTI-SOURCE LOCALIZATION

Jan Scheuing and Bin Yang

Chair of System Theory and Signal Processing
University of Stuttgart, Germany

ABSTRACT

Due to ambiguities and estimation errors, combining time differences of arrival (TDOAs) for simultaneous localization of multiple acoustic sources is a challenging task. This paper studies this problem under the framework of consistent graphs and proposes an efficient algorithm to determine TDOAs originating from the same source.

***Index Terms*—** Delay estimation, graph theory, position measurement

1. INTRODUCTION

Microphone-array based acoustic source localization systems usually consist of two estimation steps: First, the TDOA at each microphone pair is estimated using either generalized cross-correlation [1] or blind channel identification [2] methods. Knowing the positions of sensors and the velocity of sound, the source is then localized by least-squares methods like in [3, 4, 5]. This procedure has been approved in many single source scenarios. However, litte research work has been spent on the simultaneous localization of multiple sources without tracking. One problem is that each sensor pair produces more than one TDOA estimate in a multi-source scenario and it is not clear which TDOAs belong together to the same source. This paper addresses this TDOA ambiguity and proposes an algorithm to combine TDOAs of the same source.

Throughout the paper, a true (but unknown) TDOA will be denoted by $t_{a,kl,\mu\nu}$ when it results from source $a \in \{1, \dots, N\}$, sensor pair (k, l) with $k, l \in \{1, \dots, M\}$, and the corresponding paths μ and ν between the source and sensors. The indices $\mu = \nu = 0$ denote the direct paths used in localization. $\mu \geq 1$ and $\nu \geq 1$ indicate echo paths which make the localization difficult. Different TDOA estimates at sensor pair (k, l) will be represented by $\tau_{kl,\sigma}$. The problem now is to combine TDOA estimates from different sensor pairs to estimate the so called *source TDOA vectors*

$$\underline{t}_a = [t_{a,12,00},\, t_{a,13,00},\, \dots,\, t_{a,M-1\,M,00}]^T, \qquad (1)$$

each containing all direct-path TDOAs of one source a. Clearly, this combination is ambiguous, and as shown in Fig. 1, erroneous combination of TDOAs means intersection of non-matching hyperbolas for localization, which will cause phantom sources.

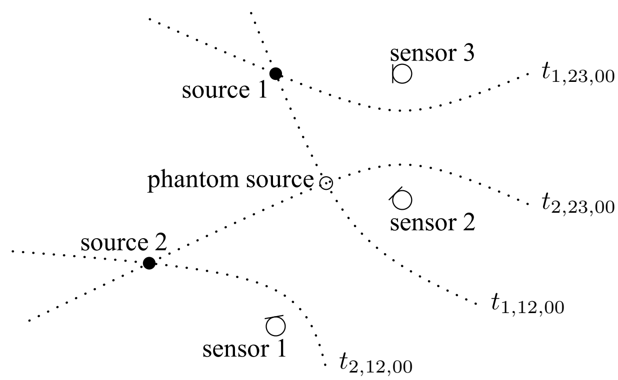

Fig. 1. Hyperbolas used for localization of $N = 2$ sources.

As proposed by the authors in [6], the ambiguity can be partly resolved by exploiting the condition that any cyclic sum of TDOAs must disappear. This means

$$t_{a,kl,\mu_k\nu_l} + t_{a,lm,\mu_l\nu_m} + \dots + t_{a,pq,\mu_p\nu_q} + t_{a,qk,\mu_q\nu_k} = 0, \quad (2)$$

where all involved TDOAs stem from the same source a and share the same paths $\mu_s = \nu_s$ with $s \in \{k, l, m, \dots, p, q\}$.

In the following, the problem of combining TDOAs is studied under the framework of *consistent graphs*. In section 2, consistent graphs are introduced and the computational complexity of different consistency checks is analyzed. Section 3 discusses some practical issues on consistent graphs of TDOA estimates. An efficient algorithm for the synthesis of consistent graphs is shown in section 4. Finally, section 5 presents some results of a real-time multi-speaker localization system using the proposed approach.

2. CONSISTENT GRAPHS

As shown in Fig. 2, the content of the source TDOA vector $\underline{t}_a$ in (1) can be visualized by a directed and labeled graph. Each node represents a sensor. Each directed branch between two nodes is labeled by the corresponding TDOA value.

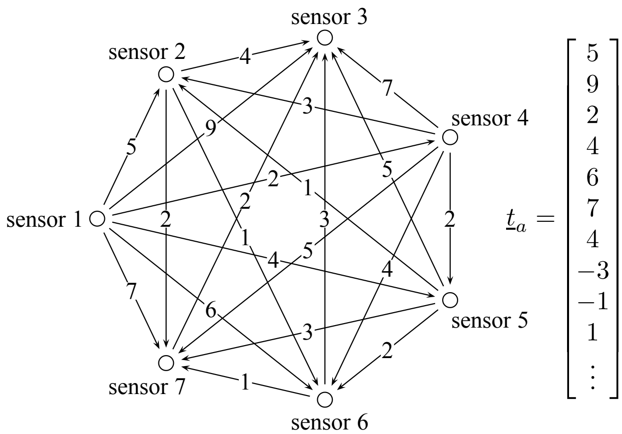

Fig. 2. A fully linked, consistent TDOA graph with 7 nodes and the related source TDOA vector $\underline{t}_a$.

The graph is called *consistent*, because the sum of all branch labels along any closed path in the graph is zero according to (2). This is very similar to Kirchhoff's second law, valid for electrical circuits (voltage graphs), except that we replace voltage by TDOA values.

In the following, different strategies to check the consistency of a graph are discussed and compaired in terms of their computational complexity. Thereby, each addition and each comparison is counted as one operation and we assume a fully linked graph with M nodes.

Consistency check for node triples

A node triple consisting of 3 nodes and 3 branches requires one addition and one comparison for its consistency check. Since an M-node graph has $\binom{M}{3}$ node triples, this strategy consumes

$$C_{\text{trip}} = 2 \cdot \binom{M}{3} = \frac{M(M-1)(M-2)}{3} \qquad (3)$$

operations.

Consistency check for n-tuples

In general, analyzing all $\binom{M}{n}$ n-tuples with $n \geq 3$ will cause

$$C_{n\text{-tup}} = \frac{n-1}{2n} M(M-1)\cdots(M-n+1) \qquad (4)$$

operations, as there are $\frac{(n-1)!}{2}$ different closed paths combining each n nodes and each path causes $(n-1)$ operations.

Consistency check for pairs

In analogy to electrical voltage and potential, a *time potential* can be defined at each node representing the time of arrival with respect to a reference node of time potential 0. The consistency check for all $\binom{M-1}{2}$ branches not including the reference node can be reduced to a comparison of their labels with the corresponding potential differences. This leads to

$$C_{\text{pair}} = 2 \cdot \binom{M-1}{2} = (M-1)(M-2) \qquad (5)$$

operations. Obviously, $C_{\text{pair}} \leq C_{\text{trip}} \leq C_{n\text{-tup}}$ holds for all sensor numbers M.

3. SYNTHESIS OF CONSISTENT TDOA GRAPHS

In multi-source localization, the aim is not the analysis but the synthesis of consistent graphs starting from sets of TDOA estimates

$$\mathbb{P}_{kl} = \{\tau_{kl,1}, \tau_{kl,2}, \dots\} \qquad (6)$$

of maybe differing cardinal numbers $|\mathbb{P}_{kl}|$ for different sensor pairs (k, l). In the ideal case, $|\mathbb{P}_{kl}|$ is equal to the number of sources N and each $\tau_{kl,\sigma}$ is equal to one of the N true direct-path TDOAs $t_{a,kl,00}$. In practice, TDOA estimates might not exactly match their true values. Some true TDOAs might not be estimated at all, and some additional $\tau_{kl,\sigma}$ might be produced by echo paths or other measurement errors. Below we discuss the resultant effects on synthesis.

3.1. Synthesis complexity

As we cannot ensure that all N true TDOAs are contained in each $\mathbb{P}_{kl}$, we will usually try to increase $|\mathbb{P}_{kl}|$ above the expected number of sources N. On the other hand, it may happen that a true TDOA is not detectable at all at some sensor pairs even for large $|\mathbb{P}_{kl}|$, e.g., due to other strong sources close to those sensors. In this case, the final graph will only be partially linked. Hence, assuming a common cardinal number $|\mathbb{P}|$ for all sensor pairs, we have to take $|\mathbb{P}|+1$ possibilities for each branch into account.

A brute force synthesis algorithm would thus check all

$$C_{\text{bf}} = (|\mathbb{P}| + 1)^{\binom{M}{2}} \qquad (7)$$

possible graphs for consistency. Attempts to reduce this high complexity to an order of $|\mathbb{P}|^{(M-1)}$ by using the idea of time potentials can be abandoned, as they implicitly assume that all TDOAs involving the reference sensor have been successfully estimated for all sources, which is not the case in practice.

Therefore, the lowest order of n-tuples to check for consistency is three.

3.2. Misleading consistencies

Besides the wanted direct-path TDOA graphs, other combinations of TDOAs can form consistent graphs as well. There are two explanations for these misleading consistencies:

Consistency due to sound reflections

Condition (2) is also valid for echo-path TDOAs like

$$t_{a,kl,0\mu}+t_{a,lm,\mu 0}+t_{a,mk,00}=0 \quad (\mu>0). \tag{8}$$

Typically, sensor l is close to a wall here. Modelling sound propagation by acoustic rays like the image source method [7], a reflecting wall has the same effect on a sensor signal as a corresponding mirrored sensor, see Fig. 3. Clearly, both the direct-path graph (Fig. 3b) and the echo-path graph (Fig. 3c) are consistent and cannot be distinguished by condition (2).

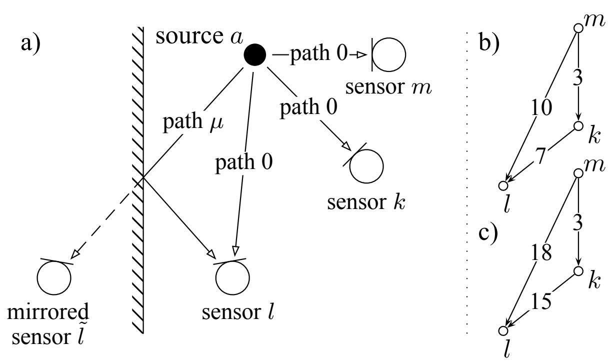

Fig. 3. A typical scenario with sound reflections where TDOA ambiguity occurs: Both path 0 and μ to sensor l produce TDOAs which combine to consistent graphs.

Accidental consistency

Equation (2) is necessary but not sufficient for TDOAs originating from a common source. Scenarios are possible, where TDOAs of different sources a, b, and c satisfy

$$t_{a,kl,00}+t_{b,lm,00}+t_{c,mk,00}=0. \tag{9}$$

For randomly distributed sources however, the probability of fully linked and accidentally consistent graphs is quite small.

3.3. Approximate consistency

Since TDOA estimates are derived from sampled and noisy sensor signals, condition (2) is only approximately fulfilled in practice:

$$|\tau_{kl,\sigma_1}+\tau_{lm,\sigma_2}+\ldots+\tau_{pq,\sigma_{\kappa-1}}+\tau_{qk,\sigma_\kappa}|<\varepsilon \tag{10}$$
$$\text{with } \tau_{kl,\sigma_1}\approx t_{a,kl,\mu_k\nu_l},\ \tau_{lm,\sigma_2}\approx t_{a,lm,\mu_l\nu_m},\ldots$$

This means that we will accept a deviation ε in the order of some sampling periods. The choice of ε depends on both the magnitude of the estimation errors $\tau_{kl,\sigma}-t_{a,kl,\mu\nu}$ and the summation length κ. In order to keep ε as small as possible, short paths are preferred for the consistency check.

4. AN EFFICIENT SYNTHESIS ALGORITHM

Due to the discussions in section 3, we choose a graph synthesis strategy based on triples. In the first step, we search for all approximately consistent TDOA triples with all sensor triples. For each sensor triple (k,l,m), let $\mathbb{T}_{klm}$ denote the set of approximately consistent TDOA triples $(\tau_{kl,\sigma},\tau_{lm,\varrho},\tau_{mk,\lambda})$ we have found. The total number of TDOA triples to be checked is

$$\sum_{k=1}^{M-2}\sum_{l=k+1}^{M-1}\sum_{m=l+1}^{M}|\mathbb{P}_{kl}||\mathbb{P}_{lm}||\mathbb{P}_{mk}|.$$

Since typically $|\mathbb{T}_{klm}|\ll|\mathbb{P}_{kl}||\mathbb{P}_{lm}||\mathbb{P}_{mk}|$, the computational complexity is significantly reduced because we only combine (approximately) consistent TDOA triples in the following.

Starting with an initial triple $(\tau_{kl,\sigma_1},\tau_{lm,\varrho_1},\tau_{mk,\lambda_1})$ from $\mathbb{T}_{klm}$ and using an additional sensor $p\in\{1,\ldots,M\}\backslash\{k,l,m\}$, we search for at least two further triples in $\mathbb{T}_{klp}$, $\mathbb{T}_{kmp}$, and $\mathbb{T}_{lmp}$ with pairwise common branch labels. If, e.g., the triples $(\tau_{kl,\sigma_2},\tau_{lp,\zeta_2},\tau_{pk,\iota_2})\in\mathbb{T}_{klp}$ and $(\tau_{lm,\varrho_3},\tau_{mp,\eta_3},\tau_{lp,\zeta_3})\in\mathbb{T}_{lmp}$ have common labels $\sigma_1=\sigma_2$, $\varrho_1=\varrho_3$, and $\zeta_2=\zeta_3$, we build a TDOA quadruple containing six different branches by combining the three triples, see Fig. 4.

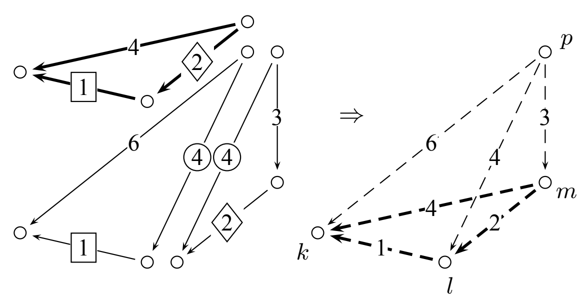

Fig. 4. Combination of an initial triple (bold) with two matching triples into one quadruple (dashed).

Two quadruples or higher order n-tuples are further combined, if they have common branches with identical labels and if at least one *branch-connecting triple* (dotted triple in Fig. 5) exists. Implicitly subjoined triples like, e.g., the triples with sensors (l,p,q) and (m,p,q) in Fig. 5 are also associated to the TDOA graph.

Continuing this procedure, we find all possible consistent and maximally linked graphs that include the chosen initial triple. Note that no further consistency check in these TDOA graphs is necessary, as each closed path is approximately consistent by construction.

After we have found all approximately consistent TDOA graphs containing the initial triple, we choose any not yet associated triple for the initialization of the next graph and repeat this procedure, until each triple is part of at least one graph. In order to reduce accidental consistency, we finally reject all graphs containing only one triple.

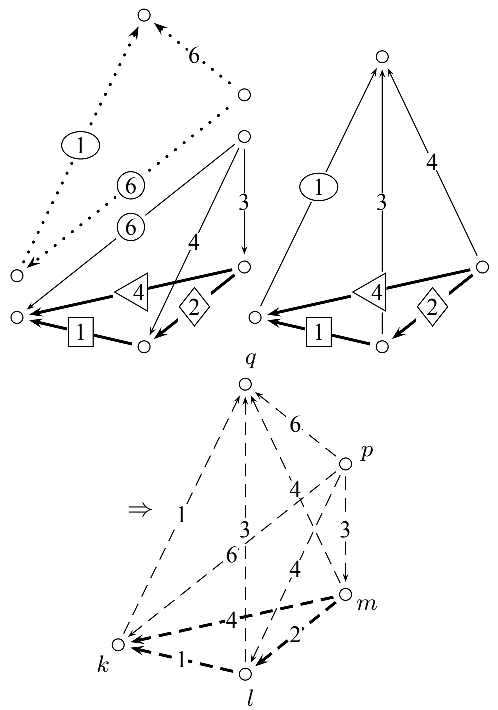

Fig. 5. Combination of two quadruples having a common initial triple (bold) with a branch-connecting triple (dotted). The result is a fully linked n-tuple of order $n = 5$ (dashed).

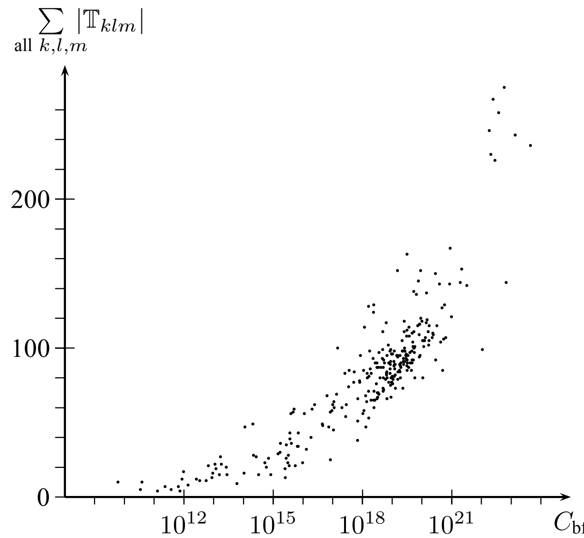

Fig. 6. Typical numbers of possible graphs versus the consistent TDOA triples resulting from the same sets of TDOA estimates.

5. EXPERIMENTAL RESULTS

Using the proposed synthesis algorithm, the combination of TDOA estimates to valid source vectors becomes a minor task for the complete acoustic source localization system in terms of computational complexity. Fig. 6 shows the number of all possible graphs versus the number of consistent TDOA triples in a real-time localization system, where two speech sources are localized by 8 microphones in a reverberant environment. The typical number of consistent TDOA triples is about 100.

Apart from echo-path graphs caused by sound reflections, the number of highly linked consistent graphs and the number of sources N match well in practice. Using the DATEMM approach in [6], we identify and reject echo paths by exploiting the autocorrelation maxima before graph synthesis. Echo-path graphs can also be determined during position estimation, as they usually lead to a larger residual error in least-squares methods.

Finally, we mention that the choice of initial triples significantly affects the speed of convergence of our synthesis algorithm. High combination rates can be achieved by starting with high-quality TDOA triples, where the cross-correlation amplitudes are large. They are represented by an internal quality measure in DATEMM.

6. REFERENCES

[1] C. Knapp and G. Carter, “The gereralized correlation method for estimation of time delay,” *IEEE Trans. on Acoustics, Speech, and Signal Processing*, vol. 24, pp. 320–327, 1976.

[2] J. Chen, Y. Huang, and J. Benesty, “An adaptive blind SIMO identification approach to joint multichannel time delay estimation,” in *ICASSP*, 2004, vol. 4, pp. 53–56.

[3] J. Delosme, M. Morf, and B. Friedlander, “Source location from time differences of arrival,” in *ICASSP*, 1980.

[4] R. Schmidt, “A new approach to geometry of range difference location,” *IEEE Trans. on Aerospace and Electronic Systems*, vol. 8, pp. 821–835, 1972.

[5] Y. Huang, J. Benesty, G. Elko, and R. Mersereau, “Real-time passive source localization: A practical linear-correction least-squares approach,” *IEEE Trans. on Speech and Audio Processing*, vol. 9, pp. 943–956, 2001.

[6] J. Scheuing and B. Yang, “Disambiguation of TDOA estimates in multi-path multi-source environments (DATEMM),” in *ICASSP*, 2006.

[7] J. Allen and D. Berkley, “Image method for efficiently simulating small-room acoustics,” *Journal of the Acoustical Society of America*, vol. 65, pp. 943–950, 1979.